

Requested Patent: JP4134523A
Title: CACHE CONTROL METHOD ;
Abstracted Patent: JP4134523 ;
Publication Date: 1992-05-08 ;
Inventor(s): OKONOGI TAKAHIRO; others: 01 ;
Applicant(s): NEC CORP ;
Application Number: JP19900256602 19900926 ;
Priority Number(s): ;
IPC Classification: G06F3/06; G06F12/08 ;
Equivalents: ;

ABSTRACT:

PURPOSE: To dynamically change capacity corresponding to the condition of use by providing a large capacity cache so as to use plural cache blocks in a prescribed size constituting the the cache in common as cache areas for all external storage devices.

CONSTITUTION: A large capacity cache 2 is interposed between a user buffer 1 and two magnetic disk devices 4 and 5 and used as the cache area for the two magnetic disk devices in common by a cache control means 3. Then, cache control is executed concerning data transfer between the user buffer 1 and the two magnetic disk devices 4 and 5. Namely, corresponding to the condition of use, the area of the large capacity cache 2 is optimumly allocated to the plural external storage devices and used for the respective external storage devices. Thus, the size of the cache area is dynamically changed, and the limited cache area is effectively utilized.

⑫ 公開特許公報(A) 平4-134523

⑬ Int. Cl.⁵G 06 F 3/06
12/08

識別記号

3 0 1 S
3 2 0

庁内整理番号

7232-5B
7232-5B

⑭ 公開 平成4年(1992)5月8日

審査請求 未請求 請求項の数 2 (全7頁)

⑮ 発明の名称 キャッシュ制御方法

⑯ 特 願 平2-256602

⑰ 出 願 平2(1990)9月26日

⑱ 発 明 者 小 此 木 崇 宏 東京都港区芝5丁目7番1号 日本電気株式会社内

⑲ 発 明 者 佐 藤 雄 一 宮城県仙台市青葉区中央4丁目6番1号 東北日本電気ソフトウェア株式会社内

⑳ 出 願 人 日本電気株式会社 東京都港区芝5丁目7番1号

㉑ 代 理 人 弁理士 境 廣 巳

明 細 書

1. 発明の名称

キャッシュ制御方法

2. 特許請求の範囲

(1) 利用者のメモリ空間上に存在するユーザバッファと外部記憶装置との間にキャッシュを介在させた計算機において、

前記キャッシュを複数のキャッシュブロックから構成される大容量キャッシュとし、該大容量キャッシュの各々のキャッシュブロックを複数の外部記憶装置用のキャッシュ領域として共通に使用してキャッシュ制御を行うことを特徴とするキャッシュ制御方法。

(2) 前記大容量キャッシュの各々のキャッシュブロックの使用頻度を管理し、一定時間ごとに、使用頻度の低いキャッシュブロックを解放して任意の外部記憶装置用のキャッシュ領域として使用可能とする請求項1記載のキャッシュ制御方法。

3. 発明の詳細な説明

(産業上の利用分野)

本発明は計算機のキャッシュ制御方法に関し、

特にユーザバッファと外部記憶装置との間に設けられたキャッシュの制御方法に関する。

(従来の技術)

外部記憶装置たとえば磁気ディスク装置は、入出力動作に機械的な動きを伴うため動作速度が遅い。このため、計算機利用者(例えばユーザプロセス)のメモリ空間上に存在し実際にデータが更新、作成、参照されるユーザバッファと磁気ディスク装置との間で直接にデータの転送を行うことはデータ入出力に要する時間の長大化を招来する。そこで、一部の計算機においては、ユーザバッファと磁気ディスク装置との間に半導体メモリ等の高速な記憶素子で構成されたキャッシュを介在させている。

このようなキャッシュを使用した計算機では、磁気ディスク装置からユーザバッファへデータを転送する際、磁気ディスク装置から読み込んだデータをキャッシュに保存(キャッシング)してからユーザバッファに転送し、再度同様に磁気ディ

スク装置からユーザバッファへのデータ転送の要求が発生した場合、そのデータが以前に転送されたことがありキャッシュに残っていると判断（キャッシュヒット）すると、そのキャッシュ内のデータをユーザバッファへ転送することで、実際に磁気ディスク装置から転送するのに比べてより短い時間でデータの入力が行えるようにしている。

逆に、ユーザバッファから磁気ディスク装置への転送時には、ユーザバッファのデータをキャッシュに書き込んだ時点で転送要求元に対し出力の完了を通知することでデータ出力にかかる見掛け上の応答時間を短縮している。なお、キャッシュから磁気ディスク装置への書き込みは、その後、転送要求元とは非同期に行われる。

そして、計算機が複数の磁気ディスク装置を有する場合、従来は、各磁気ディスク装置毎に専用のキャッシュを設けている。

（発明が解決しようとする課題）

ところで、多くの入出力要求が特定の磁気ディスク装置に集中して頻繁に発生すると、その磁気

用してキャッシュ制御を行うようにしている。

また、前記大容量キャッシュの各々のキャッシュブロックの使用頻度を管理し、一定時間ごとに、使用頻度の低いキャッシュブロックを解放して任意の外部記憶装置用のキャッシュ領域として使用可能としている。

（作用）

外部記憶装置の台数を n 台とし、1 台当たり従来において容量 P のキャッシュを割り当てていたとすると、本発明では例えば容量 $P \times n$ の 1 つの大容量キャッシュを設け、それを構成する所定サイズの複数のキャッシュブロックを全ての外部記憶装置用のキャッシュ領域として共通に使用する。共通に使用するので、極端な場合、或る 1 台の外部記憶装置のみに入出力要求が発生し、残りの外部記憶装置に入出力要求が発生していないとすると、入出力要求の発生している外部記憶装置が容量 $P \times n$ の大容量キャッシュ全体を使用することになり、実質的にその磁気ディスク装置のキャッシュ領域が拡大したことになる。また、各外部記

ディスク装置用のキャッシュの空き領域が不足し、キャッシュによる効果が薄れてくるが、従来においては、使用できるキャッシュの容量はその磁気ディスク装置に予め割り当てられたキャッシュの容量で規定され、それ以上動的に拡大することはできなかった。このため、キャッシュ効果の低下を有効に防止することができなかった。

そこで、本発明の目的は、個々の磁気ディスク装置用に使用できるキャッシュ領域の実質的な容量を使用状況に応じて動的に変化させることができるキャッシュ制御方法を提供することにある。

（課題を解決するための手段）

本発明は上記の目的を達成するために、

利用者のメモリ空間上に存在するユーザバッファと外部記憶装置との間にキャッシュを介在させた計算機において、

前記キャッシュを複数のキャッシュブロックから構成される大容量キャッシュとし、この大容量キャッシュの各々のキャッシュブロックを複数の外部記憶装置用のキャッシュ領域として共通に使

用装置に対し平均的に入出力要求が発生している場合は、各外部記憶装置に対しほぼ容量 P のキャッシュ領域が与えられることになり、その状態で入出力要求頻度が低下した外部記憶装置が現れると、それが使用していたキャッシュブロックが残りの外部記憶装置用のキャッシュ領域として使用される。このように、使用頻度に応じて各外部記憶装置用のキャッシュ領域が動的に拡大、縮小する。

（実施例）

次に、本発明の実施例について図面を参照して詳細に説明する。

第 1 図は本発明の一実施例の構成図であり、ユーザバッファ 1 と 2 台の磁気ディスク装置 4、5 との間に大容量キャッシュ 2 を介在させ、キャッシュ制御手段 3 によって、大容量キャッシュ 2 を 2 台の磁気ディスク装置用のキャッシュ領域として共通に使用して、ユーザバッファ 1 と 2 台の磁気ディスク装置 4、5 との間におけるデータ転送にかかるキャッシュ制御を行うようにしたもので

ある。ここで、ユーザバッファ 1 は、利用者のメモリ空間上に存在し、実際にデータが更新、作成、参照されるバッファであり、例えば計算機の主記憶上に設けられている。また、大容量キャッシュ 2 は、所定サイズの複数のキャッシュブロック CB 1 ~ CB n で構成されており、例えば拡張記憶装置上に設けられる。また、磁気ディスク装置 4、5 は同じ条件のものであり、各々その上にファイルが作成されている。なお、説明の便宜上、ファイルを構成するブロックのサイズとキャッシュブロックのサイズとは等しいものとし、入出力要求はブロック単位で行われるものとする。

第 2 図はキャッシュ制御手段 3 がキャッシュ制御に使用する制御テーブルの構成例を示す。この制御テーブル 3 0 は、大容量キャッシュ 2 のキャッシュブロック CB 1 ~ CB n と 1 対 1 に対応するエントリ E 1 ~ E n で構成され、各エントリに、キャッシュブロック識別子 3 1、使用フラグ 3 2、使用頻度カウンタ 3 3、反映フラグ 3 4、格納データ識別子 3 5 および格納先磁気ディスク情報 3

6 が格納される。ここで、キャッシュブロック識別子 3 1 は文字通り各々のキャッシュブロックを識別するための情報、使用フラグ 3 2 は当該キャッシュブロックが使用中か否かを示すフラグ、使用頻度カウンタ 3 3 は当該キャッシュブロックの使用頻度に応じた値を示すカウンタである。例えば、当該キャッシュブロックに新たなデータが格納された時点で所定値を示し、その後、一定時間経過毎に 1 だけ減算され、当該キャッシュブロックのデータが参照、更新のためにアクセスされる毎に任意の値たとえば 2 だけ加算されるようなカウンタである。また、反映フラグ 3 4 は当該キャッシュブロックのデータが磁気ディスク装置に反映されているか否かを示すフラグ、格納データ識別子 3 5 は当該キャッシュブロックに格納されているデータの識別子、格納先磁気ディスク情報 3 6 は当該キャッシュブロックに格納されているデータをどの磁気ディスク装置のどの場所に格納するかを示す情報である。

第 3 図乃至第 5 図はキャッシュ制御手段 3 の処

理例を示し、以下、各図を参照して本実施例の動作を説明する。

第 1 図において、図示しないユーザプロセス等から例えば磁気ディスク装置 4 上に作成された或るファイルの或るブロックの入力要求が出された場合、キャッシュ制御手段 3 は第 3 図に示す処理を行う。先ず、キャッシュ制御手段 3 は第 2 図の制御テーブル 3 0 を参照して、入力要求されたファイルのブロック（要求データ）が大容量キャッシュ 2 上に存在するか否かを調べる（S 1）。

存在した場合、すなわち入力要求されたファイルのブロックに対応する格納データ識別子 3 5 を持つエントリが大容量キャッシュ 2 に存在した場合は、そのエントリに対応するキャッシュブロックに格納されているデータを読み出してユーザバッファ 1 へ転送する（S 2）。このとき、そのエントリの使用頻度カウンタ 3 3 を 2 増加させる。

他方、存在しない場合は、磁気ディスク装置 4 からファイルのブロックを大容量キャッシュ 2 に読み込む必要があるため、空きのキャッシュブ

ロックが存在するか否かを第 2 図の制御テーブル 3 0 を参照して調べる（S 3）。存在する場合は、その空きのキャッシュブロックに磁気ディスク装置 4 から読み出したファイルのブロックを格納すると同時にユーザバッファ 1 へ転送する（S 4）。このとき、使用したキャッシュブロックに対応する第 2 図の制御テーブル 3 0 のエントリに必要な情報を設定する。即ち、例えばキャッシュブロック識別子 CB 1 のキャッシュブロックにファイルのブロックを格納したとすると、エントリ E 1 の使用フラグ 3 2 に使用中、使用頻度カウンタ 3 3 に所定値、反映フラグ 3 4 に反映済、格納データ識別子 3 5 に今回読み出したファイルのブロックの識別子、格納先磁気ディスク情報 3 6 に今回読み出したファイルのブロックが磁気ディスク装置 4 上のものであり且つそのどの場所に存在したかを示す情報を、それぞれ設定する。また、処理 S 3 で空きのキャッシュブロックが存在しないと判断したときは、磁気ディスク装置 4 から該当するファイルのブロックを読み出し、大容量キャッシ

ュ 2 を経由せずに直接にユーザバッファ 1 に伝送する (S 5)。

次に、第 1 図において、図示しないユーザプロセス等がユーザバッファ 1 上に読み込んだ磁気ディスク装置 4 の或るファイルの或るブロックを更新し、その出力要求を出した場合、キャッシュ制御手段 3 は第 4 図に示す処理を行う。先ず、キャッシュ制御手段 3 は第 2 図の制御テーブル 30 を参照して、出力要求されたファイルのブロックの更新前のブロック (更新前のデータ) が大容量キャッシュ 2 上に存在するか否かを調べる (S 11)。

存在した場合、すなわち出力要求されたファイルのブロックに対応する格納データ識別子 35 を持つエントリが大容量キャッシュ 2 に存在した場合は、そのエントリに対応するキャッシュブロックに格納されているデータを出力要求されたユーザバッファ 1 上のデータで置き換える (S 12)。このとき、そのエントリの使用頻度カウンタ 33 を所定値に設定すると共に反映済フラグ 34 を未反映に設定する。また、ユーザプロセス等に対し

て出力完了を通知する。

他方、存在しない場合は、大容量キャッシュ 2 にキャッシングする必要があるため、空きのキャッシュブロックが存在するか否かを第 2 図の制御テーブル 30 を参照して調べる (S 13)。存在する場合は、その空きのキャッシュブロックにユーザバッファ 1 上のデータを置き込む (S 14)。このとき、使用したキャッシュブロックに対応する第 2 図の制御テーブル 30 のエントリに必要な情報を設定する。即ち、例えばキャッシュブロック識別子 CB 2 のキャッシュブロックにファイルのブロックを格納したとすると、エントリ E 2 の使用フラグ 32 に使用中、使用頻度カウンタ 33 に所定値、反映フラグ 34 に未反映、格納データ識別子 35 のファイルのブロックの識別子、格納先磁気ディスク情報 36 に今回のファイルのブロックがどの磁気ディスク装置 4 のどの場所に格納すべきかを示す情報を、それぞれ設定する。また、出力要求元に対しては大容量キャッシュ 2 への置き込み終了時に出力完了を通知する。また、処理

S 13 で空きのキャッシュブロックが存在しないと判断したときは、ユーザバッファ 1 上のファイルのブロックを大容量キャッシュ 2 を経由せずに直接に磁気ディスク装置 4 に伝送する (S 15)。

ユーザプロセス等からの入力要求時および出力要求時には、上述したような動作が行われるが、それらと非同期に遅延置き込み動作とキャッシュブロック解放動作とが行われる。

遅延置き込みは、キャッシュ制御手段 3 がユーザプロセスの出力要求とは非同期に、第 2 図の制御テーブル 30 を参照して、使用フラグ 32 が使用中で且つ反映フラグ 34 が未反映になっているエントリの有無を調べ、そのようなエントリが存在するときは、そのエントリの格納先磁気ディスク情報 36 が示す磁気ディスク装置 4 または 5 の格納場所へ、同エントリに対応するキャッシュブロックに格納されたデータを置き込み、その後同エントリの反映フラグ 34 を反映済に設定することで行われる。

また、キャッシュブロックの解放は、キャッシ

ュ制御手段 3 が第 5 図の処理を定期的に行うことで行われる。先ず、キャッシュ制御手段 3 は、内部ポインタ P に第 2 図の制御テーブル 30 の先頭エントリを指示する値を設定する (S 21)。次に、内部ポインタ P が指し示す制御テーブル 30 のエントリの使用フラグ 32 が使用中か否かを調べ (S 22)、使用中でないなら処理 S 28 へ進み、使用中であればそのエントリの使用頻度カウンタ 33 の値が所定値 CN 以下か否かを調べる (S 23)。使用頻度カウンタ 33 の値が所定値 CN 以下の場合は、同エントリに対応するキャッシュブロックを参照、更新の為に最近あまりアクセスされていないことになるので、そのエントリの反映フラグ 34 を調べ (S 24)、それが未反映を示していればそのキャッシュブロックの内容を該当する磁気ディスク装置 4 に書き込んだ後 (S 25)、また反映済フラグ 34 が反映済を示していればそのような動作を省略して、同エントリに対応するキャッシュブロックを解放する (S 26)。即ち、同エントリの使用フラグ 32 を未使用に設

定する。また、使用頻度カウンタ 33 が所定値 C N より大きいときは、そのエントリの使用頻度カウンタ 33 の値を 1 だけ減じ (S 27)、処理 S 28 へ進む。

処理 S 28 では、内部ポインタ P の値が第 2 図の制御テーブル 30 の最終エントリ E n に相当する値 M A X 以上となったか否かを判定し、未だそれに達していない場合は、内部ポインタ P の値を 1 エントリ分進め (S 29)、処理 S 22 に戻って残りのエントリに対し上述したと同様の処理を繰り返す。また、内部ポインタ P の値が M A X 以上になっていれば、第 5 図の処理を終了する。

以上本発明の実施例について説明したが、本発明は以上の実施例にのみ限定されず、その他各種の付加変更が可能である。例えば、キャッシュ制御手段 3 による制御方法として、第 2 図のような制御テーブルを使用する以外に、各磁気ディスク装置毎に割当中のキャッシュブロックの使用状況をヘッダ等のリストで管理し、一定時間毎に使用頻度の低いキャッシュブロックを検出して共通の

空きプールで管理する等の他の任意の方法を採用することが可能である。

(発明の効果)

以上説明したように、本発明によれば、使用状況に応じて大容量キャッシュの領域を複数の外部記憶装置に対して最適に割り当てて各々の外部記憶装置用に使用できるキャッシュ領域のサイズを動的に変化させることができるため、限られたキャッシュ領域の有効利用が促進され、計算機全体としての入出力処理の応答時間を短縮することが可能となる。

4. 図面の簡単な説明

第 1 図は本発明の一実施例の構成図、

第 2 図はキャッシュ制御手段 3 が使用する制御テーブルの構成図、

第 3 図はキャッシュ制御手段 3 が行う入力要求処理の一例を示す流れ図、

第 4 図はキャッシュ制御手段 3 が行う出力要求処理の一例を示す流れ図および、

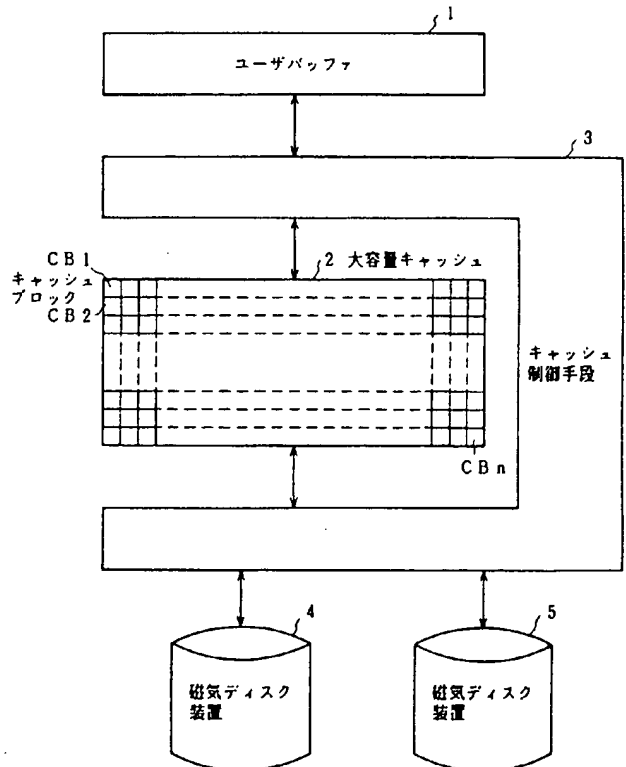
第 5 図はキャッシュ制御手段 3 が行うキャッシ

ュブロック解放処理の一例を示す流れ図である。

図において、

- 1 … ユーザバッファ
- 2 … 大容量キャッシュ
- 3 … キャッシュ制御手段
- 4, 5 … 磁気ディスク装置
- C B 1 ~ C B n … キャッシュブロック

特許出願人 日本電気株式会社
代理人 弁理士 境 廣 巳



本発明の一実施例の構成

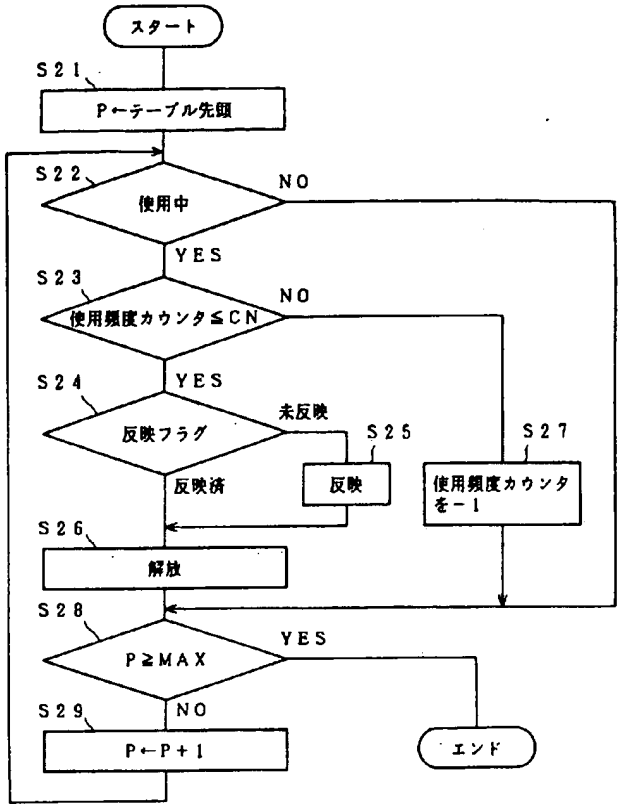
第 1 図

30

	31 キャッシュ ブロック識 別子	32 使用 フラグ	33 使用頻度 カウンタ	34 反映 フラグ	35 格納データ 識別子	36 格納先磁気 ディスク情 報
E1	CB1					
E2	CB2					
	⋮	⋮	⋮	⋮	⋮	⋮
En	CBn					

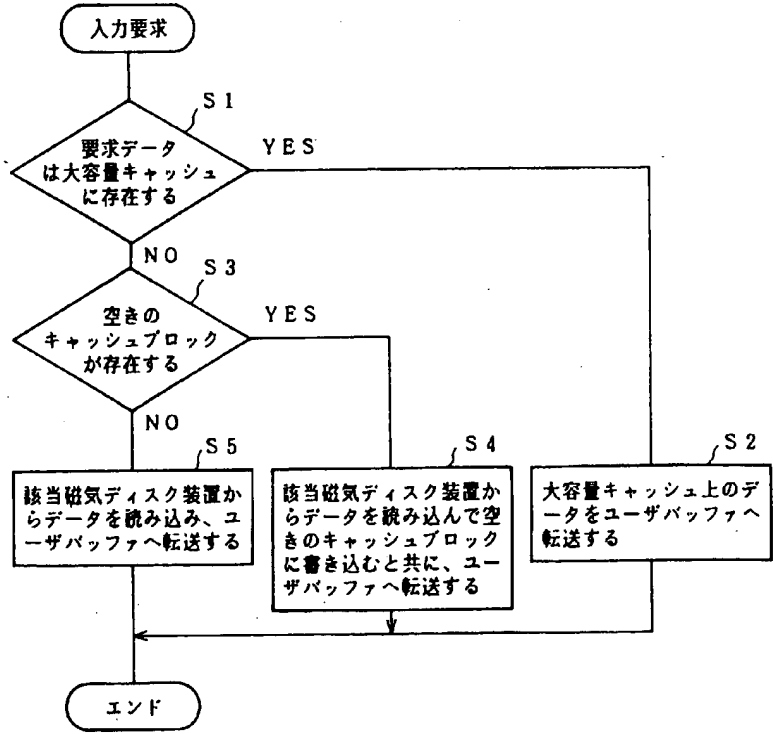
キャッシュ制御手段3が使用する制御テーブルの構成

第 2 図



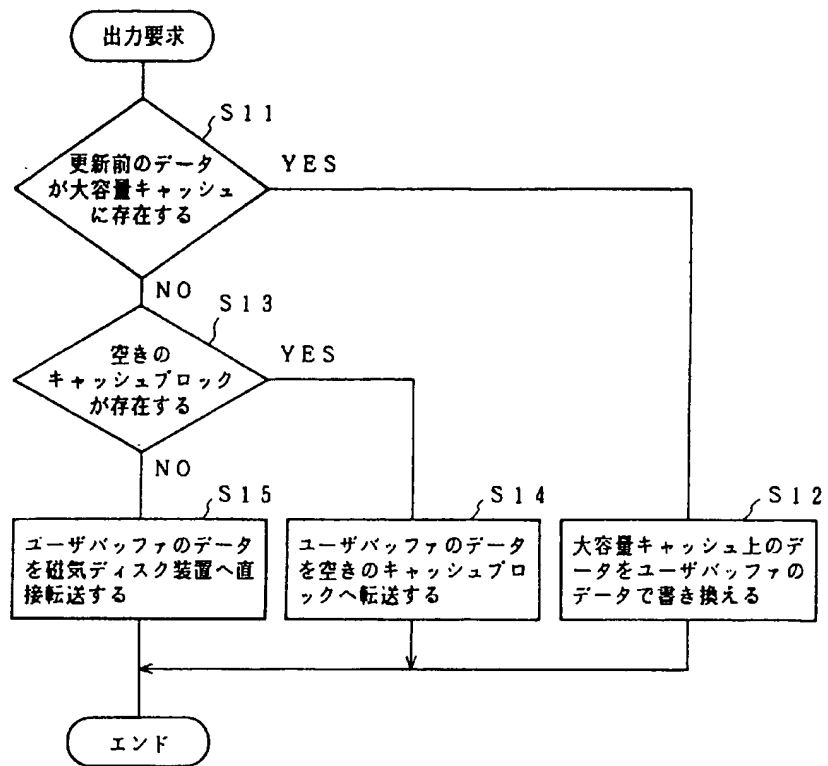
キャッシュ制御手段3が行うキャッシュブロック解放処理例

第 5 図



キャッシュ制御手段3が行う入力要求処理例

第 3 図



キャッシュ制御手段3が行う出力要求処理例

第 4 図